Sindh Univ. Res. Jour. (Sci. Ser.) Vol.50(001)101-106 (2018)

http://doi.org/10.26692/sujo/2018.01.0017

SINDH UNIVERSITY RESEARCH JOURNAL (SCIENCE SERIES)

Comparative Analysis of Spatial Interpolation Techniques for Mapping Annual Mean Rainfall Estimation within a Mountainous Region

A. N. LAGHARI, H.U. ABBASI*

Energy and Environment Engineering Department, Quaid – e – Awam University Nawabshah

Received 2nd April 2017 and Revised 26th January 2018

Abstract: The complex topography, poor gauge representativity and uneven density make it an uphill task to accurately map precipitation in mountainous regions. This challenge was confronted with the evaluation of four different mapping techniques: Inverse Distance Weighting (IDW), Ordinary Kriging (OK), Spline and Regression Kriging (RK). An evaluation of the resulting georasters using 1) cross-validation statistics, 2) a spatial cross-consistency test and 3) a water balance analysis reveals that the techniques ignoring the information on co-variables yield the largest prediction errors. Mean error and root-mean-square error values suggested that the most biased methods were IDW and spline, with a bias almost 2 to 5 times higher than ordinary kriging. The best model accounted for mean precipitation analysis is regression Kriging, with a mean error and root mean square error values of 1.38 mm and 72.36 mm respectively, which represents 42 % less bias and 16 % higher accuracy than OK results. Comparative performances show that the regression analysis made it possible to judiciously evaluate the variable patterns and get fairly accurate values at un-gauged locations where geographical information compensated the poor availability of local data.

Keywords: Spatial interpolation, precipitation, Digital elevation model, topography

1. INTRODUCTION

Climate change; almost every one now are agreed that climate is undergoing significant changes (Laghari *et al*. 2012); much attention being paid to analyze the effect of climate changes on hydrology and water resource management. Distributed hydrological models are one of the sources gaining enormous importance in featuring and examining overall impacts in mountainous environment (Viviroli *et al* 2009). The continuous regular spaced estimates of climatic variables are prerequisite input requirements for proper functioning of spatially distributed hydrological models (Laghari *et al*. 2012). Precipitation is the one and most critical input parameter in hydrological modeling. However, this input is subject to uncertainty, as a result of measurement errors, systematic error in interpolation techniques and stochastic error due to the random nature of rainfall (Buytaert *et al*, 2006). An accurate spatial estimate of precipitation is the key to the performance of the above models. Even a small bias resulting from the used interpolation method can drastically affect the conclusions, which are better addressed with a detailed knowledge of the spatial distribution patterns (Beven, 2001a).

This challenge increases many folds in mountainous environments, where topography is complex with enormous influence over variables; gauging stations are sparse and concentrated in the valleys. Gauging stations on higher elevations are typically poorly represented. The acquirement of accurate information for the mountainous range is therefore an uphill task, all this making it difficult to model pattern analyses. In such cases; when no single method is optimal nor superiority of a specific interpolation method has been established, the performance depends on the variable under study, spatial configuration, and the assumptions used (Creutin and Obled, 1982; Weber and Englund, 1992, 1994; and Prudhomme *et al*. 1999). To obtain the best representative precipitation mapping technique for this particular mountainous region, it is essential to compare the results by applying alternative methods to the same data set in rugged terrain. To achieve the above objective, this study examines a variety of stochastic and deterministic mapping methods to estimate the values at un-gauged locations.

2. DATA AND MODELS

2.1. The study area

The study area of the greater Kitzbühel region located in the eastern part of the Austrian Province of Tyrol covers an area of about 2000 km^2. A detailed description of the study area can be found in (Vanham *et al*., 2008). The basin has strong seasonal rainfall patterns and rugged topography with altitudes ranging from 400 m to 2400 m above mean sea level. Available mean annual precipitation time series of 30 years (1961-1990) from 14 stations are used for the spatial

++Corresponding Author E-mail: a.n.laghari@quest.edu.pk mashaalnasirlaghari@gmail.com

*Center for Environmental Sciences University of Sindh Jamshoro

precipitation interpolation for the study area **(fig.1).** The resolution of all spatial georasters is 250m.

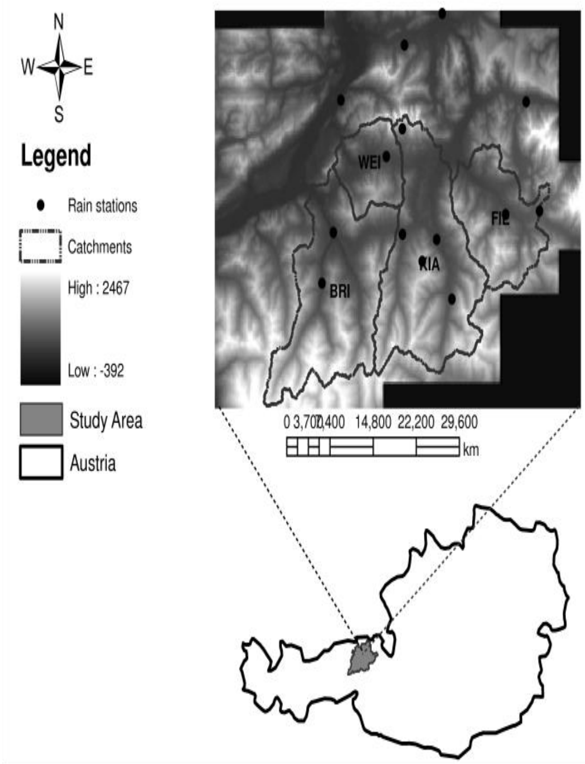

Fig. 1: Terrain map of the area under study (resolution of 250 m) with locations of 14 rain gauge stations

2.2. Spatial interpolation techniques

The mapping methods used here can broadly be categorized into two approaches: deterministic and stochastic. The deterministic approaches used for precipitation interpolation are inverse distance weighting (IDW) and spline. Ordinary kriging (OK) is based on stochastic approach while regression kriging (RK) centers on both - the combination of linear regression and its kriged residuals (Odeh *et al*., 1994, 1995; Knotters *et al*. 1995). The regression is based on the spatial correlation of predicted variable and the co-variables (Moore *et al*., 1993; Richardson and Edmonds, 1987; Chaplot *et al*., 2000a; Thompson *et al*., 2001). The model is assumed to be able to remove topographical drifts through regression and the intrinsic hypothesis still assumed valid and can produce high stationary field (Holdaway, 1996; and Prudhomme *et al*. 1999). The approach with different denominations is successfully used for climatic data mapping (Philips *et al*. 1992; Martinez-Cob 1996; Odeh *et al*. 1995; and Holdaway 1996). All techniques are briefly discussed here. The detailed description of algorithms can be referred to (Goovaerts, 1997; Hengl *et al*., 2003; Odeh *et al*., 1995).

2.3. Variogram modeling

The descriptive statistics of the average annual precipitation measurement data of all 14 weather stations showed that precipitation values were log-normally distributed **(Table-1)**. Subsequently, log-transformation and a multiplied factor of 1000 for avoiding any numerical error during kriging were applied (Martinez-Cob, 1996).

Table 1: Descriptive statistics of study area

	DEM (m asl)	PREP (mm)	LPREP (mm)
Minimum	500	1206	7.09
Mean	1264	1483	7.28
Maximum	2427	2484	7.81
Std. Deviation	387	332	0.19
Median	1230	1363	7.22
Skewness	0.4	2.031	1.50
Kurtosis	2.2	6.947	5.14

DEM-Digital elevation model, PREP-Precipitation, LPREP-log precipitation

The geostatistical analyst of Arc-GIS 9.2 was used to compute sample direct-semivariogram of precipitation and residuals of linear regression of precipitation on geographical variables. After visual inspection of the sample semivariograms, a Gaussian model was chosen as best fit model for direct-semivariogram. Model parameters **(Table 2)** and validity were checked through a trial and error procedure until satisfactory cross validation statistics were achieved.

Table 2: Model parameters of direct semivariogram of precipitation (LPREP) and their residuals of linear regression of LPREP on geographical variables

Mapping techniques	Model Parameters		
	Nugget[1]	Sill[1]	Range (km)
LPREP	653	3127	43.560
Residual of LPREP	601	2080	28.535

[1] LPREP and residuals of linear regression of LPREP on geographical variables $\log (mm)^{2} * 10^{-6}$

The theoretical model was fitted to the data and the anisotropy was checked in all directions. Up to a distance of 10 km, no deviations in variogram were observed in any direction. However, between about 10 to 24 km, in (N-S) & (E-W) orientation, variograms are similar to the isotropic variogram. The same orientation starts increasing variance above 24 km and again follows the isotropic behavior from 32 km. At distances above 15 km, (NE-SW) orientation shows deviation than for isotropy till 32 km, while the inter-site variance in (NW-SE) orientation also does not become stable, with a nugget-sill value (700-2930) reached at 44 km. The little fluctuation in variation of (E-W), (N-S) orientation clearly indicates that there are no big differences in annual rainfall estimates in these directions; however in NW-SE, and NE-SW directions, maximum difference of rainfall were observed between 27 to 48 km.

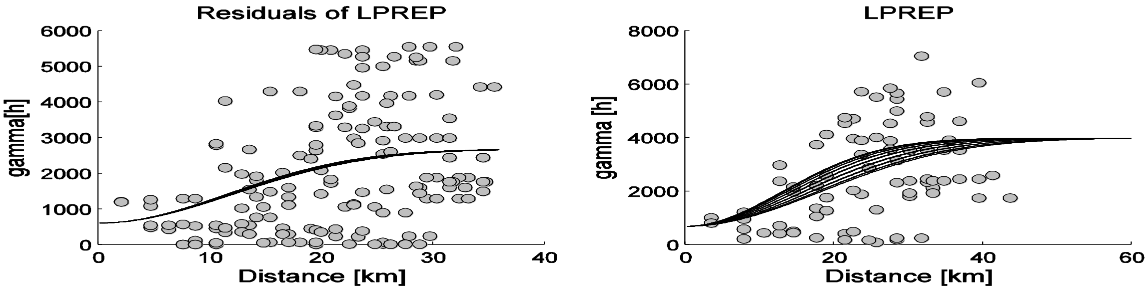

Fig.2. Gaussian variograms of LPREP and residuals of LPREP for different orientations

The above analysis confirms that the stationary hypothesis does not hold for the whole region, but only locally. These trends violet the assumption of stationarity on which this ordinary kriging is based. Despite these evidences of non stationarity in the data, isotropy was considered, and an isotropic model was used in the final mapping of ordinary kriging with maximum neighborhood of interpolation of 27 km. The same procedure was adopted for residual kriging. After trend removal through multiple linear regression of LPREP on geographical variables, overall behavior of variogram fluctuation matches with isotropic variogram. The anisotropic trial observations confirmed that after trend removal, the variability is reduced at great extent and closely matched to isotropic model in all directions. Finally the isotropic model of residuals was used for precipitation mapping. The model parameters are listed in **(Table-2)**. The estimated kriged precipitation values were back-transformed before final estimates of precipitation.

3. PREDICTION PERFORMANCE

The performance of mapping techniques was assessed by means of three strategies: 1) by cross-validation statistics, 2) by spatial cross-consistency and 3) by a water balance approach. The cross-validation procedure is based on removing one sample location (measurement station) from the data set at a time and calculating the value of the removed sample with remaining data points. This routine was followed for each measurement station. The comparative indices were then used as a measure of prediction quality by mean prediction error (ME) and root mean square error (RMSE), which are defined as follows:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\hat{T}_i - T_i)$$

$$RMSE = \left\{\frac{1}{n}\sum_{i=1}^{n}(\hat{T}_i - T_i)^2\right\}^{0.5}$$

Where n = Number of validation points,

$\hat{T}_i$ & T_i = Predicted and observed values at location i.

The ME criterion is used to check the conditional bias property, while the RMSE criterion assesses the precision quality. A smaller value of RMSE indicates higher accuracy while the higher value indicates vice versa. Cross validation statistics can be used to find optimal mapping technique, however, the presence of short range correlations in data may rise questions regarding the reliability of its statistical results (Hutchinson 1998 a).

As second step to evaluate the reliability and consistency of predictions, spatial cross-consistency approach was adopted (Hofierka *et al*. 2002). All statistical parameters of different calculated precipitation mapping estimates were compared with a referenced precipitation map (RPM). This referenced precipitation map was carefully produced during a 4 year project (www.waterpool.org) in which different experts from different institutes were involved, and results were consistent with water balance estimates.

Finally all calculated precipitation maps including the referenced precipitation map were evaluated by means of a general water balance approach.

Q (Discharge) = P (precipitation) – ET (evapotranspiration) – ΔS (storage changes)

Gridded discharges were calculated for each mapping technique as a result of subtraction of actual evapotranspiration grid estimates from interpolated mapping precipitation grid estimates. The actual evapotranspiration values are obtained from the Hydrological Atlas of Austria. Storage changes can be ignored, as for long-range mean annual water balances; it was assumed that there is no sensible change in the water contents of different reservoirs, e.g. groundwater, snow cover (Hofierka *et al*. 2002; Kuhn & Escher-Vetter 2004). The difference between the calculated discharges with observed discharges gives a measure for the reliability and consistency of the precipitation.

4. RESULTS AND DISCUSSION

Mean error and root-mean-square error values suggested that the most biased methods were IDW and spline, with a bias almost 2 to 5 times higher than ordinary kriging (Table 3). According to mean error results, regression kriging (RK) is the best interpolation technique, producing about 42% less bias than ordinary kriging. RMSE results indicate that all four methods produce higher uncertainty in predicted values. However, regression kriging (RK) produced comparatively the least uncertainty. Thus there is clearly a significant improvement in the estimation of performance with taking co-variables into account (RMSE decreases from 111.3 to 72). Due to the limited number of 14 stations, it is difficult to create a subset of stations for adopting a commonly used validation strategy.

Table 3: Cross validation statistics of four mapping techniques applied in study region: mean estimation error (ME), root-mean-square error (RMSE).

Mapping techniques	Cross-Validation Statistics	
	ME	RMSE
IDW	**-14.24**	**111.3**
Spline	**7.54**	**98.7**
OK	**2.39**	**86.38**
RK	**1.38**	**72.36**

Therefore, the results of the computed precipitation maps were assessed through spatial cross-consistency with a referenced raster map (RPM). The Custer et al 1996 [38] adopted the similar approach for assessment of the cross-consistency between computed techniques and referenced precipitation maps.

Table 4: Summary statistics of percent differences between referenced precipitations map (RPM) and computed precipitation estimates by IDW, Spline, OK and RK for entire region. Negative values indicate a lower value with respect to the referenced map value, while positive values indicate vice versa.

Interpolation Techniques		IDW	Spline	OK	RK
Total Min.		9.00	-34.27	8.50	3.00
Total Max		-14.31	10.89	-10.80	-19.60
Std. deviation		-37.00	37.00	-14.00	3.00
Variance		-60.00	86.00	-25.00	7.00
Percentile	25	2.12	-6.00	-9.00	3.36
Percentile	50	-12.25	-7.26	-16.60	4.92
Percentile	70	-14.88	-9.84	-18.40	6.30
Total Mean		-10.00	-8.00	-12.00	4.00

The results of the comparison between all 4 techniques and the referenced precipitation map (RPM) are presented in Table 4. The maximum visible distinction among interpolation techniques are statistical parameters of minimum, mean, standard deviation, variance and maximum. Although the RK method yielded the high range of maximum difference, the all other lowest statistical parameters indicate that the RK method gives the most promising results. The above results clearly show that the incorporation of topographical information has increased the range of maximum values. However, it has significantly helped in producing the matching estimates of standard deviation, variance range, mean and percentile ranges. The percentile differences show that the RK precipitation map is closer to the referenced map than all other technique estimates. The IDW and OK estimates are within the range of original sample values because only sample values were used for interpolation. The same distribution pattern follows for minimum and maximum values, whereas minimum and maximum values obtained for RK were generally outside the original sample value ranges, due to the incorporation of topographical information. To assess the performances of interpolation techniques at more detailed level (Table 5); we evaluated performances at a) four basins and b) the whole region divided in four elevation zones. The location of stations is highly biased; i.e. 86 % of the stations are located below 1000 m, an elevation zone only representing 31.5 % of the total area. Zones 2 & 3 each contain 39 % and 26.21% area respectively with one station each, whereas zone 4 contains 3.17 % area without any station.

Table 5: Comparative statistics of percent differences between referenced precipitations map (RPM) and computed precipitation estimates by IDW, Spline, OK and RK at four basins & elevation zones. Negative values indicate the less value from the referenced map and positive values indicate vice versa.

Basins			INTERPOLATION TECHNIQUES			
			IDW	Spline	OK	RK
Weißache (1 stations)			-3.28	-31.30	-8.04	-5.83
Fieberbrunner Ache (2 stations)			-9.52	-4.47	-9.52	3.93
Brixenthaler Ache (2 stations)			-15.93	-15.23	-19.66	2.65
Kitzbüheler Ache (5 stations)			-4.60	4.33	-6.77	9.0
elevatio wise	Zone-i	Below 1000 meter	1. 28	1.07	-0.50	-2.01
	Zone-ii	1000 - 1500 meter	-8.41	-8.16	-23.4	4.85
	Zone-iii	1500 - 2000 meter	-18.29	-13.69	-21.15	9.99
	Zone-iv	Over 2000 meter	-30.49	-25.35	-33.46	5.34

The mean annual values at four basin levels almost follows the same pattern of mean annual values of whole region. A basin-wise result clearly shows the superiority of RK method in three basins, while in basin four, Kitzbüheler Ache; all the other three techniques produced the better estimates. The better estimates of RK over OK, at one side endorse the improvement in results after trend removal, while on other side justifies RK application in sparse network high altitude regions. The results of Kitzbüheler Ache, which contain directly five stations and 3 other stations just surrounding boundary also strengthen the many other study's conclusions; the superiority of conventional techniques i.e. IDW, spline over geostatistical techniques in areas of concentrated network stations. The above conclusions were further cemented in altitude-wise analysis; where the techniques without elevational information, IDW, spline and OK comparatively performed well below 1000 m, the zone comprising 86% of all stations. The same time all these three techniques proved their ineffectiveness at higher elevational zones, where the increasing consistency in under estimation between all these three techniques reaches from 8% at zone-2 to 33% at higher altitude zone-4, clearly marginalize their abilities in higher sparse data zones.

The relative poor performance of OK at higher zones was also realized during semivariogram modeling, where it was observed that the spatial dependence was much higher when stations in zone 1 were only considered and weakened when station in zone 2 and 3 included one by one respectively, but due to comparative analyses of mapping techniques and importance of each station, we included all stations in final mapping. The increasing good results of RK over other technique at respective higher zones also proves the usefulness of geographical information in mapping estimates at higher sparse data zones, where co-variable values were almost the only ones used for interpolation. In other words, RK technique provides the excellent recipe of precipitation and topographical features that can be realized in RK map (figure 3), where regional structural information was incorporated through linear regression, while OK accounted for precipitation patterns. The structured contrast can also be seen in RK and other techniques produced maps.

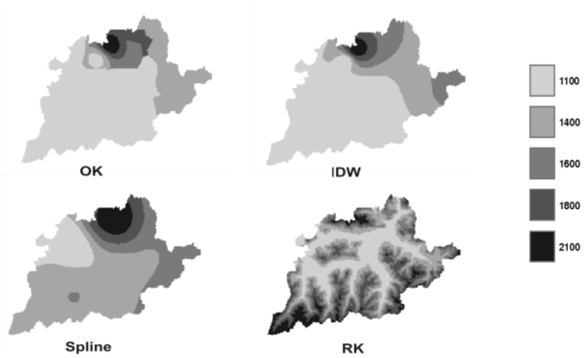

Fig.3. The mean annual precipitation maps developed through interpolation of 14 gauge station values using Ordinary Kriging (OK), (IDW), SP, and (RK).

To further verify the authenticity of the referenced precipitation map (RPM) and its usefulness for the comparative analyses of the four different interpolated precipitation maps, a water balance approach was adopted. Mean annual evapotranspiration estimates of whole region were taken from hydrological atlas of Austria. The runoff was calculated by subtracting the evapotranspiration estimates from each computing precipitation map estimates. The percent difference between computed runoff for each mapping technique and gauged runoff clearly proved the validity of referenced precipitation map results with just 0.4% difference, followed by RK runoff difference with 2.7%. The runoff difference produced within the other three techniques ranged from 13% to 20%. These findings through the water balance approach pacify the cross-validation and cross-consistency results, i.e. the overall superiority of the RK technique in this high altitude region with sparse data.

5. CONCLUSION

The wide-ranging evaluation demonstrates that the simple, easy to use RK technique offers reliable and reasonably accurate estimates of mean annual precipitation. The superiority of RK were observed in high altitude sparse data zones, especially proves the usefulness of geographical information in areas of unobtainable mapping variable, while on the other side, comparative good results of conventional techniques i.e. IDW and spline over geostatistical techniques also justifies their applications in low elevational concentrated zone of network stations.

REFERENCES:

Ahmed, S., G. de Marsily, (1987). Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specic capacity. Water Resources Research, 23(9):1717{1737

Beven, K., (2001). How far can we go in distributed hydrological modelling? Hydrology and Earth System Sciences 5, 1–12.

Buytaert, W., R. Celleri, P. Willems, B. De Bièvre, G. Wyseure, (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. Journal of Hydrology, 329 (3-4): 413-421

Creutin, J. D., C. Obled, (1982). Objective analysis and mapping techniques for rainfall fields: An objective comparison. Water Resour. Res. 18(2), 413-431

Chaplot, V., C. Walter, P. Curmi, (2000). Improving soil hydromorphy prediction according to DEM resolution and available pedological data. Geoderma , 97, S. 405-422.

Hengl, T., G. Heuvelink, A. Stein, (2003). Comparison of kriging with external drift and regression kriging. Technical report. International Institute for Geo-information Science and Earth Observation (ITC) Enschede.

Knotters, M., D. J. Brus, J. H. O. Voshaar, (1995). A comparison of kriging, cokriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. Geoderma, 67: 227-246.

Laghari, A. N., D. Vanham, W. Rauch (2012): To What extend does climate change result in a shift in Alpine hydrology? A case study of Austrian Alps. Hydrological Sciences Journal, 57:1, 103-117

Legates, D. R., C. J. Willmott, (1990). Mean seasonal and spatial variability in global surface air temperature. Theoretical Application in Climatology 41: 11–21.

Matheron, G., (1969). Le krigeage universel, Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, 1

Moore, I. D., P. E. Gessler, G. A. Nielsen, G. A. Peterson, (1993). Soil attribute prediction using terrain analysis. Soil Science Society of America Journal, 57, S. 443-52.

Odeh, I. O. A., A. B. McBratney, D. J. Chittleborough, (1994). Spatial prediction of soil properties from land form attributes derived from a digital elevation model. Geoderma, 63, 197-215.

Odeh, I. O. A., A. B. McBratney, D. J. Chittleborough, (1995). Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma, 67, 215-226.

Prudhomme, C., D. W. Reed, (1999). Mapping extreme rainfall in a mountainous region using geostatistical techniques: A case study in Scotland, International Journal of Climatology 19, 1337-1356.

Richardson, J. L., W. J. Edmonds, (1987). Linear regression estimations of Jenny`s relative effectiveness of state factors equation. Soil Science, 144(3), 203-208.

Stallings C, S. Khorram, Z. Guo, (1992). Linking Gleams and GIS ASAE Paper 92–3613. American Society of Agricultural Engineers: Nashville, TN.

Vanham D., E. Fleischhacker, W. Rauch, (2008). Technical Note: Seasonality in alpine water resources management – a regional assessment. Hydrology and Earth System Sciences , 12 (1), 91-100.

Viviroli, D., M. Zappa, J. Gurtz, R. weingartner, (2009). An introduction to the hydrological modelling system PREVAH and its pre- and post- processing tools. Environmental Modelling and Software, 24, 1209-1222.

Weber, D. D., E. J. Englund, (1992). Evaluation and comparison of spatial interpolators. Mathematical Geology, 24: 381–391.

Weber, D. D., E. J. Englund, (1994). Evaluation and comparison of spatial interpolation II. Mathematical Geology, 26: 589-603.

Webster, R., M. Oliver, (2001). Geostatistics for Environmental Scientists. John Wiley & Sons, UK.